DOCUMENT RESUME

ED 273 093 FL 015 833

AUTHOR Nunan, David
TITLE The Ghost in the Machine: An Examination of the Concept of Language Proficiency.
PUB DATE 86
NOTE 15p.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Journal Articles (080)
JOURNAL CIT Prospect; v2 n2 1986

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Communicative Competence (Languages); Comparative Analysis; Definitions; Language Processing; *Language Proficiency; Language Research; *Language Tests; Linguistic Theory; *Rating Scales; Research Needs; *Second Language Learning

ABSTRACT

Proficiency, which refers to a concept that is more complex than is generally acknowledged, is an overworked term in second language teaching that lacks a satisfactory operational definition. Linguistic knowledge is currently defined most often in terms of what an individual is able to do with the knowledge, as in the movement for competency-based education. The simple idea of proficiency as the ability to perform real-world tasks with a specified degree of skill becomes problematic when the issue of proficiency assessment arises. All proficiency scales assess a mixture of factors from diverse domains. Most of the performance factors have not been empirically validated to determine whether learners really do act in the ways described by the scales or to what degree the skills mastered in one domain are transferable to another. The only performance factors subjected to empirical validation have been syntax and morphology, and it has been found that some aspects of proficiency descriptions conflict with what learners are able to do at other stages. Research into the divisibility of the construct of general proficiency has not been as fruitful as originally anticipated. Despite some valuable research, efforts at producing an operational definition of proficiency seem to be circular and will not be resolved without more empirical study. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED273093

The Ghost in the Machine: An Examination of the Concept of Language Proficiency.

David Nunan,
National Curriculum Resource Centre,
Adelaide.

PROSPECT, 2, 2, 1986

Abstract

'Proficiency' is an overworked term in language teaching. Despite this, there are no satisfactory operational definitions of the concept. This paper questions some of the assumptions underlying the use of the concept, and argues that the development of our understanding and appreciation of the concept of proficiency is more likely to result from empirical research than theoretical speculation.

Introduction

> When we describe people as exercising qualities of mind, we are not referring to occult episodes of which their overt acts and utterances are effects; we are referring to those overt acts and utterances themselves. There are, of course, differences ... between describing an action as performed absent-mindedly and describing a physiologically similar action as done on purpose, with care or with cunning. But such differences of description do not consist in the absence or presence of an implicit reference to some shadow-action covertly prefacing the overt action. They consist, on the contrary, in the absence or presence of certain sorts of testable explanatory-cum-predictive assertions.
>
> (Ryle 1949:26)

There has been so much talk of late of the concept of proficiency that it is time the concept were subjected to critical analysis. Most language programs subscribe to the concept in one form or other, yet the term itself is rarely explicated. This failure to define proficiency is one reason why language curricula have suffered from arrested development, and also why the concept itself is controversial. Curricular activities such as setting objectives, grading content and testing and evaluation are difficult enough to carry out at the best of times. When there is uncertainty and confusion about what it is that is being taught, the tasks become even more difficult.

In this paper, I should like to undertake a conceptual analysis of the concept 'proficiency'. In the course of this analysis it will be necessary to examine the related concepts 'competence', 'performance',

FL 015 833

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
David Nunan
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

and 'achievement'. These terms are also relatively ill-defined in the literature.

While the paper does not pretend to be the last word on the subject, it does attempt to articulate the major points of contention, and thereby circumscribe the outer parameters of the concept of proficiency. It is hoped that, thus circumscribed, there may be greater appreciation of the complexity of a concept which seems to be taken so much for granted. It is also hoped that the paper will stimulate a productive debate on the nature of language proficiency.

Concepts in Contention

Linguists, both theoretical and applied, seem obsessed with conceptual universes in which creatures come in pairs. Thus, we have 'langue' and 'parole', 'competence' and 'performance', 'use' and 'usage', 'form' and 'meaning', 'context' and 'cotext', 'cohesion' and 'coherence'. In was Chomsky (1965) who gave prominence to the competence/performance distinction (although the theoretical distinction between the terms was not Chomsky's). For Chomsky, 'competence' refers to mastery of the principles governing language behaviour. 'Performance' refers to the manifestation of these internalised rules in actual language use. The terms have come to be used to refer to what a person knows about a language (competence) and what a person does (performance). More recently, the term 'communicative competence' has come into currency. This refers to knowledge of the rules of use and appropriacy and includes linguistic competence. This might seem reasonably straightforward. Unfortunately, there are a number of complicating factors. To begin with, there does not seem to be anything like universal agreement on what is meant by 'knowing'. Does 'knowing the rules of language' mean being able to recite them? If so, most native speakers must be classed as incompetent. According to Chomsky, however, native speakers are, by definition, competent. Diller (1978) suggests that:

> Linguists are sometimes hesitant to say that ordinary people "know" the rules of their language, because lingusts themselves have such a hard time trying to formulate these rules explicitly.
>
> (p.26)

He points out that children can create phonological rules for nonsense words through a process of analogy, although they are unable to give a formulation for these rules. He goes on to ask:

> But if children are not able to formulate the rules of grammar which they use, in what sense can we say that they "know" these rules? This is the question which has bothered linguists. The answer is that they know the rules in a functional way, in a way which relates the changes in abstract grammatical structure to changes in

> meaning. Knowledge does not always have to be formulated. Children can use tools before they learn the names for these tools.
>
> (p.26-27)

For Diller, then, knowledge need not be conscious but may manifest itself in the ability to use the language. However, this would seem to render the competence/performance distinction rather uncertain.

Krashen (1981, 1982) further confuses the issue by suggesting that knowledge of linguistic rules is the outward manifestation of one psychological construct (learning), while the manifestation of these rules in use is the manifestation of another construct (acquisition). Rea (1985) has since questioned the need for a 'competence' construct by suggesting that, as we can only observe instances of performance, not competence, the competence/performance distinction is redundant. She brings this view into line with communicative language teaching by proposing yet another bifurcation; communicative performance and non-communicative performance. In her scheme of things, what you see is what you get.

It would seem, therefore, that we have reached a point where linguistic knowledge is to be defined in terms of what an individual is able to do with that knowledge. This is reinforced by a recent movement in ESL in the United States; competency-based ESL. As though there were not enough confusion over terminology, this movement is using 'competence' to refer to things learners can do with language; that is, it is used in roughly the same sense as 'performance' in the earlier competence/performace distinction. The concept of competency-based education (CBE) has been borrowed into ESL from the field of adult education where it is used to specify the skills needed by adults to function in today's society in areas such a communication, computation, problem solving and interpersonal relationships.

In ESL, "a competency is a task-oriented goal written in terms of behavioural objectives" (CAL 1983:9). The following characteristics of CBE as it relates to ESL have been articulated:

1. Teaching ESL to competencies requires the instructional focus to be on functional competencies and life coping skills. It is not what the students know about language but what they do with the language.

2. Assessment is built in. Once the competency has been identified, it also serves as a means of evaluating student performance. Since it is performance based, assessment rests on whether the student can perform the competency or not. The only problem is to establish the level at which the student can perform the competency.

3. Competencies are based on an assessment of student needs.

(CAL 1983:11-13)

Within the literature, some writers use the term 'proficiency' as an alternative to 'competency' (see, for example, Higgs 1982). Richards, on the other hand, makes a clear distinction between 'competence' and 'proficiency', although he characterises the concept of proficiency in the same way as CBE characterises competency. This can be seen in the following quote:

> 1. When we speak of proficiency, we are not referring to knowledge of a language, that is, to abstract, mental and unobservable abilities. We are referring to performance, or, that is, to observable or measurable behaviour... Whereas competence refers to what we know about the rules of use and the rules of speaking of a language, proficiency refers to how well we can use such rules in communication.
>
> 2. Proficiency is defined with reference to specific situations, settings, purposes, settings and tasks
>
> 3. Proficiency also implies the notion of a skill. It refers to the degree of skill with which a learner can perform a task.
>
> 4. Lastly, proficiency refers to the integration and application of a number of subskills in performing particular tasks.
>
> (Richards 1985:3-4)

The Assessment of Proficiency.

The foregoing discussion demonstrates the confusion surrounding a number of key concepts relating to the nature of language. This confusion is due partly to the inconsistent application of terms to concepts and partly to the confusion over the nature of the concepts themselves.

If the Richards' line is followed, proficiency, simply put, refers to the ability to perform real world tasks with a specified degree of skill. This might seem straightfoward enough. However, problems arise when we turn to the issue of proficiency assessment. This is because the psychological reality of the construct 'proficiency'.

A 'construct' is an abstract psychological quality which is assumed to underly various forms of observable human behaviour. Thus, in the domain of intelligence testing, the construct 'verbal reasoning' is postulated to account for performance on certain types of test. Such constructs are derived by examining patterns of intercorrelations between test scores. Thus, if high correlations are recorded between scores on Tests A, B and C, it is assumed that a single underlying mental ability is being tapped by the tests.

What has this to do with language proficiency? I shall try to demonstrate that, in the case of proficiency rating scales, the construct 'general language proficiency' has not been empirically derived, but has been assumed to exist because the notion is intuitively appealing. Having assumed the existence of the construct, language testers have created rating scales to measure it in language learners. In other words, they have reversed the procedure adopted by researchers in the field of intelligence testing.

What exactly is it that the creators of proficiency scales are trying to test? We can get some idea from an examination of the instruments themselves. One such instrument is the American Council on the Teaching of Foreign Languages Provisional Proficiency Guidelines. Here is the generic description for Intermediate-High level speaking:

Able to satisfy most survival needs and limited social demands. Shows some spontaneity in language prodiction by fluency is very uneven. Can initiate and sustain a general conversation but has little understanding of the social conventions of conversation. Developing flexibitity in a range of circumstances beyond immediate survival needs. Limited vocabulary range necessitates much hesitation and circumlocution. The commoner tense form occur but errors are frequent in formation and selection. Can use most question forms. While some word order is established, errors still occur in more complex patterns. Cannot sustain coherent structures in longer utterances or unfamiliar situations. Ability to describe and give precise information is limited. Aware of basic cohesive features such as pronouns and verb inflections, but many are unreliable, especially if less immediate in reference. Extended discourse is largely a series of short, discrete utterances. Articulation is comprehensible to native speakers used to dealing with foreigners, and can combine most phonemes with reasonable comprehensibility, but still has difficulty in prodicing certain sounds in certain positions or in certain combinations, ans speech will usually be laboured. Still has to repeat utterances frequently to be understood by the general public. Able to produce some narration in either past or future.

(Cited in Savignon and Berns 1984:228-9)

In introducing a modified version of this scale (the ASLPR), Ingram defines proficiency as the mobilisation of linguistic knowledge to carry out communication tasks. Statements of proficiency must therefore be made in behavioural terms. According to Ingram, the scale is designed to measure a construct he calls 'general proficiency'. Such a construct is defined and defended in the following way:

> ... language occurs only in situations, and, if the proficiency descriptions are related to particular situations, one could be accused of measuring only proficiency in specific situations i.e. one would not be measuring general proficiency, but proficiency in specific registers. On the other hand, language varies from situation to situation; it varies according to who is using it, to

> whom and about what subject ... in other words, it wouls seem as though one cannot speak of general proficiency so much as proficiency in a language in this situation or that, in this register or that. Yet such a view would seem to be counter-intuitive. If we say that X speaks Chinese ... we do not mean that X can only give a lecture on engineering in Chinese ... Rather, when we say that someone can speak a language, we mean that that person can speak the language in the sorts of situations people commonly encounter. That is, there are certain everyday situations in which we, as human beings living in a physical and social world, are necessarily involved ... General proficiency, then, refers to the ability to use the language in these everyday, non-specialist situations.
>
> (Ingram 1984:10-11)

Ingram is postulating the existence of an underlying, unobservable construct called 'general proficiency' which, because it is unobservable, must be inferred from learners' performance in specific situations. However, it is not to be confused with the ability to perform in specified situations (i.e., it is more than an achievement test): "the ASLPR seek to measure the underlying general proficiency rather than the fulfilment of an absolutely specified task in an absolutely specified situation" (Ingram 1984:11). Learners must therefore be given the opportunity of performing in situations and contexts with which they are familiar.

What precisely is it that we are assessing in making our rating of a given learner? We know that it is not specific situational or contextual knowledge, so presumably this lets out lexical knowledge and the ability to discuss certain topics or themes. What is left when these are taken away are global, impressionistic judgements of the learner's current stage of development in a number of areas including morphological and syntactic development, fluency, pronunciation, sociocultural knowledge, mastery of discourse and so on. All proficiency scales have this same mixture of factors from diverse domains. Carroll (1981), for instance, lists size, complexity, range speed, flexibility, accuracy, appropriacy, independence, repetition and hesitation. These are generally all rendered down into a single numerical index or descriptor such as '1+' or 'Novice - Low'. The 'general language ability' underlying proficiency scales look suspiciously like Oller's (1979) unitary competence construct, which is dealt with in the next section.

Ingram defines 'general proficiency' in terms of the ability of the learner to carry out tasks in 'certain everyday situations in which we, as human beings living in a physical and a social world are necessarily involved ...' General proficiency, then refers to the ability to use the language in these everyday, non-specialist situations. However, it could be argued that this ability does not necessarily represent an ability which all language leanrers have but is, in fact, another register.

Another difficulty with proficiency scales is related to the means whereby learners are assessed. This is generally through some form of oral interview. However, it is difficult to see how such interviews can allow one to make valid judgements about the learner's ability to carry out real world tasks.

Let us look in greater detail at the sample description taken from the ACTFL scale. According to the scale, the learner, at 'Intermediate - High' level is able to satisfy survival needs and limited social needs. At this level, the learner's performance will be characterised by the following features:

- can satisfy some survival needs and limited social demands
- shows some spontaneity
- fluency is very uneven
- can initiate and sustain a general conversation
- has little understanding of the social conventions of discourse
- has limited vocabulary range
- commoner tense forms occur, but errors are frequent in formation and selection
- can use most question forms
- basic word order is established
- errors occur in more complex patterns
- cannot sustain coherent structures in longer utterances
- has limited ability to describe and give precise information
- is aware of basic cohesive feature
- extended discourse is largely a series of short, discrete utterances
- articulation is comprehensible to native speakers used to dealing with foreigners
- can combine most phonemes with reasonable comprehensibility
- has difficulty in producing certain sounds in certain positions or in certain combinations
- speech willusually be laboured
- has to repeat utterances frequently to be understood by the general public
- can produce some narration in either past or future

The vague, impressionistic speech features are indicators of the postulated 'general proficiency', the ghost in the machine, which is driving the user's communication skills. These features manifest themselves in tasks such as the following:

- coping with less routine situations in shops, post office, bank (e.g. asking for a larger size, returning an unsatisfactory purchase) and on public transport (e.g. asking passenger where to get off for unfamiliar destination)

- explaining some personal symptoms to a doctor

- communicating routine needs and basic details of unpedictable occurrences

(Ingram 1984:43)

The use of such scales is fraught with hidden dangers. The scales themselves tend to take on ontological status (that is, there is a tendency to assume that such a construct as "survival proficiency" really exists, rather than being something constructed to account for observable features of learners' speech). The scales themselves have not been empirically validated to determine if learners really do act in the ways described by the scales. Nor have the task types been validated. For example, is the task of 'returning an unsatisfactory purchase' of the same order of difficulty as 'explaining some personal symptoms to a doctor' as suggested by the scale? Do the two tasks draw on the same linguistic and communicative resources? How do we test for these things? As already pointed out, most ratings are conducted through an oral interview. While these will provide us with information on the learner's ability to take part in interviews, they are unlikely to tell us much about the learner's ability to 'explain personal symptoms to a doctor', or to 'ask passenger where to get off for unfamiliar destination'.

To summarise, then, proficiency refers to the ability to perform certain communicative tsks with a certain degree of skill. Degree of skill will be determined by mastery of a complex set of enabling skills which will include syntax and morphology, fluency, socio-cultural knowledge, phonology, and discourse. Whether or not these can or should be taught, or whether they will emerge spontaneously as a function of learning to perform certain communicative acts is a hotly debated issue within the profession.

A popular means of assessing proficiency is the use of rating scales. A subjective and impressionistic assessment of the learner's current 'general proficiency' level is made through an oral interview. Level of proficiency is set by such performance factors as fluency, mastery of syntax and so on. These are assumed to correlate with the ability to perform real world tasks. However, the link between performance factors and task difficulty has never been empirically validated. In addition, the degree to which skills mastered in one domain transfer to another is open to dispute. We must assume that some transfer occurs (otherwise there would hardly be any point in teaching).

The only performance factors to have been subjected to empirical validation are syntax and morphology. Here, it has been found that some aspects of proficiency descriptions are at adds with what learners are actually able to do at different stages (Johnson 1985).

The Divisibility of Proficiency

It would seem that the construct of general proficiency must draw a large part of its theoretical rationale from an assumption that the construct itself is a single psychological entity, in much the same way as the construct of 'verbal intelligence' is assumed to be a single psychological entity. Without this assumption, it is difficult to see

how claims about the comparability of different performance factors and task types could be made.

In fact, unlike proficiency scales, the question of whether or not a single construct underlies the ability to use language has been empirically investigated, and it is to these investigations that we now turn.

The name most commonly associated with research into the factorial structure of language proficiency is John Oller. Oller suggested that a single psychological construct underlay language proficiency. He called this construct a 'pragmatic expectancy grammar', and characterised it in the following manner:

> Language use is viewed as a process of interacting plans and hypotheses concerning the pragmatic mapping of linguistic contexts onto extralinguistic ones. Language learning is viewed as a process of developing such an expectancy system.
>
> (Oller 1979:50)

Oller put his theory to the test by utilising procedures similar to those used in intelligence testing. He analysed the scores of language learners on a wide range of tests to determine the degree of correlation between them (correlation refers to the degree to which subject scores on one test covary with scores on another test).

Oller wanted to test which of three possible hypotheses about language learning were supported by the data. These hypotheses were as follows:

Hypothesis 1 The Divisibility Hypothesis: Language proficiency consists of a number of discrete skills.

Hypothesis 2 The Indivisibility Hypothesis: Proficiency consists of a single skill.

Hypothesis 3 The Partial Divisibility Hypothesis: In addition to a general skill, common to all areas of language use, there will be skills uniquely related to various language modalities.

In order to test these hypotheses, Oller set himself the task of finding:

> testing procedures that will generate variances that are unique to tests that are supposed to measure different things. Either the indivisibility hypothesis or the partial divisibility hypothesis allows for a large general factor (or component of variance) common to all language tests. The difference between these alternatives is that the indivisibility hypothesis allows only for a general component of test variance. Once such a component is accounted for, the indivisibility hypothesis predicts that no additional reliable variance will remain to

be accounted for.

(Oller 1979:425)

By carrying out a complex set of statistical procedures, called factor analysis, on his data, Oller found that, "once the general factor predicted by the indivisibility (or unitary competence) hypothesis was extracted, essentially no meaningful variance was left in any of the tests" (p.429). The indivisibility hypothesis was thus upheld.

In non-statistical terms, the results suggested that performance on tests, say, of reading, draw on the same underlying language skill as tests of listening; or that aspects of a macroskill (such as pronunciation, fluency, control of syntax and vocabulary for 'speaking') are all part of an underlying 'proficiency'.

The implications of such a finding for language teaching were unequivocal. If all language performance derived from a single underlying psychological construct, then differentiated curricula (including needs-based courses) were redundant, and efforts to develop such courses a waste of time. The only thing needed would be a series of learning tasks which engaged the learner's interest and which were at the appropriate level of difficulty.

Oller went even further. He suggested that:

> Implications of the foregoing findings for education are sweeping. They suggest a possible need for reform that would challenge some of the most deeply seated notions of what school is about - how schools fail and how they succeed. The potential reforms that might be required if these findings can be substantiated are difficult to predict. Clearly they point us in the direction of curricula in which the focus is on the skills required to negotiate symbols rather than on the 'subject matter' in the traditional sense of the term. They point away from the medieval notion that psychology, grammar, philosophy, English, history and biology are intrinsically different subject matters. Physics and mathematics may not be as reasonably distinct from English literature and sociology as the structure of universities implies.
>
> (Oller 1979:457)

As it turned out, follow-up research did not substantiate Oller's findings. In 1983, Volmer and Sang were able to demonstrate that, on statistical grounds alone, Oller's results were suspect. Since then, research such as that reported by Bachman and Mack (1986) suggest that proficiency consists of a number of factors which are related to each other in extremely complex ways. At present, the consensus seems to be that proficiency is a multidimensional construct. Brindley (1986) suggests that the unitary/divisibility debate:

> has now been substantially resolved in favour

> of a multidimensional view, allowing, however, for the existence of a weaker general factor than was originally postulated.
>
> (Brindley 1986:ii)

Future Directions

Unfortunately, then, Oller's seductively simple characterisation of language proficiency could simply not be sustained. In fact, we have not progressed very far in our efforts to produce an operational definition of proficiency. While there is value in ongoing theorising and conceptualisation, if we are not to continue going around in circles, it is critically important that we put the whole area on more secure empirical foundations.

Proficiency rating scales are a case in point. Such scales can be derived in one of two ways. They can be derived through intuition, whereby teachers, applied linguists and curriculum specialists sit down and produce descriptions based on what they imagine learners should be able to do at various stages. Alternatively, the scales might derived empirically, that is, by carrying out studies of what learners are actually able to do at various stages. This second alternative will be a lengthy and complicated process, and those interested in the development or refinement of proficiency scales might be tempted to follow an intuitive course.

Unfortunately, intuition does not seem to be a very reliable guide, as is evidenced by a number of studies. Johnston (1985), for instance, has demonstrated that there is often a marked discrepancy between what teachers and coursebook writers believe learners should be able to do at various stages, and what they can actually do. Alderson (1986), found major disagreements between applied linguists as to the difficulty level of reading comprehension questions. On those questions about which the experts did agree, there was little correlation between their judgements and the actually difficulty as measured by the ability of the learners to answer the questions.

One possible option for teaching institutions would be to avoid the problem altogether by developing instruments to measure achievement rather than proficiency. This would enable the proficiency issue to be neatly sidestepped. The measures themselves would also have high face validity (assuming they were properly constructed) because they would be directly assessing what a given curriculum had set out to teach.

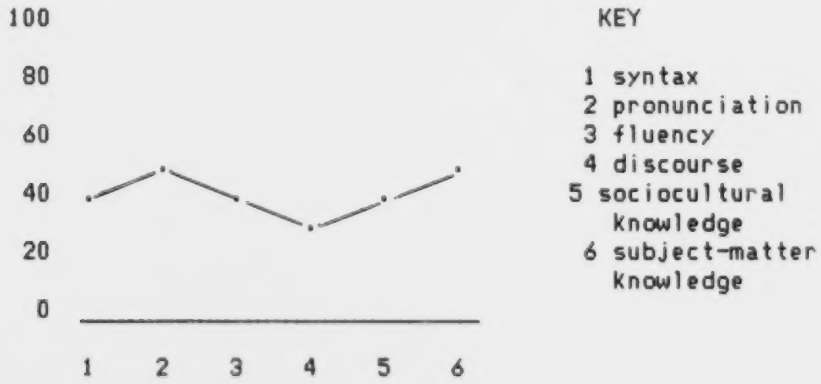
There has been talk of late about the development of 'proficiency profiles', as an alternative to the relatively imprecise proficiency rating scales (Campbell 1985). However, unless such profiles are empirically derived, they are unlikely to be any more satisfactory than the rating scales they are intended to replace. From the discussion in the preceeding section, we know that proficiency consists of a plurality of factors, which interact in complex ways. Profiles would need to

identify and articulate these diverse factors and give some indication of how they interrelated in real communication.

At the very least, it is likely that proficiency profiles would need to be constructed along two dimensions. One dimension would be what might be called 'subject factors' (i.e. factors relating to the learner) and would include such things as syntactic and morphological mastery, pronunciation, fluency, sociocultural knowledge, lexical knowledge, subject matter knowledge and so on. The other dimension could be termed 'task factors' and these would be located within the task itself. They would include degree of contextual support, cognitive demand, amount of assistance provided, psycholinguistic processing difficulty, and degree of stress.

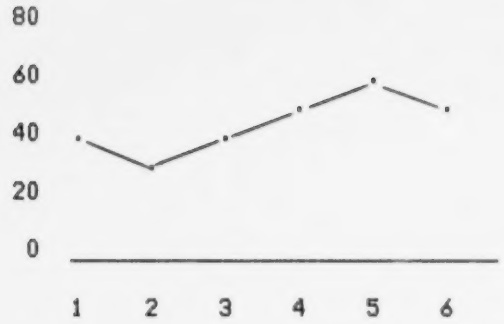
If the intention of those advocating the development of profiles intend that they be stable indices of the learner's current state of development, then there is bound to be disappointment. Given the complex interaction between 'subject' and 'task' factors, the profiles themselves are likely to be highly unstable. Thus, the 'subject profile' for a given learner might look quite different from one task to another, according to the degree to which 'task' factors influence the ability of the learner to carry out the task.

In order to clarify this point, let us consider an example. Let us assume that reasonably precise instruments for measuring the mobilisation of the subject factors of syntactic mastery, pronunciation, fluency, discourse, sociocultural knowledge and subject-matter knowledge have been developed and empirically validated. The instruments are then used to measure the performance of Subject A on Task X, which results in the following profile:



On Task Y, however, in which the task factors such as processing difficulty and contextual support are different, Subject A may have the following profile.

100



The hypothesis suggested here, then, is that proficiency profiles must take into consideration two sets of factors which will interact in complex ways to make the profile for any given learner unstable. This hypothesis needs to be empirically tested to estimate the degree of instability. The variability of different learner profiles also need to be computed. If there is reasonable correlation in the direction and magnitude of variability across tasks, generalisations can start to be made.

One way of controlling for variability might be to allow learners to nominate the task types on which they are to be assessed. However, it could well be claimed that, if learners are only rated against tasks they have been taught, the resultant profile would be one of achievement rather than proficiency.

CONCLUSION

This paper has explored some of the complex issues surrounding the concept of 'language proficiency'. The purpose of the paper is to demonstrate just how complex these issues are. It is certainly not the intention of the paper to denigrate the work of those who have made many valuable contributions, particularly in the field of proficiency assessment, but to indicate areas where empirical research is needed to support the conceptual work which has already been undertaken. It is suggested that our attempts at providing an operational definition are still at a primitive stage, and that a great deal of empirical investigation needs to be carried out in order to take us beyond our current stage of knowledge.

REFERENCES

Bachman, L. and Mack, M. 1986. A causal analysis of learner characteristics and second-language proficiency. Paper presented at the TESOL 20th Annual Convention, Anaheim, California, March 1986.

Brindley, G. 1986. The Assessment of Second Language Proficiency: Issues and Approaches. Adelaide: National Curriculum Resource Centre.

Campbell, J. 1985. Towards Active Voice. Canberra: Department of Immigration and Ethnic Affairs.

Carroll, B. 1981. Testing Communicative Performance. Oxford: Pergamon.

CAL, 1983. From the Classroom to the Workplace: Teaching ESL to Adults. Washington DC: Center for Applied Linguistics.

Chomsky, N. 1965. Aspects of the Theory of Syntax. Cambridge Mass.: M.I.T. Press.

Diller, C. 1978. The Language Teaching Controversy. Rowley Mass.: Newbury House.

Higgs, T. (ed.) 1984. Teaching for Proficiency: The Organising Principle. Lincolnwood Ill.: National Textbook Company.

Ingram, D. 1984. Australian Second Language Proficiency Ratings. Canberra: Australian Government Publishing Service.

Johnston, M. 1985. Syntactic and Morphological Progressions in Learner English. Canberra: Department of Immigration and Ethnic Affairs.

Krashen, S. 1981. Second Language Acquisition and Second Language Learning. Oxford: Pergamon.

Krashen, S. 1982. Principles and Practice in Second Language Acquisition. Oxford: Pergamon.

Lee, Y., C. Fok, R. Lord and G. Low (eds.). 1985. New Directions in Language Testing. Oxford: Pergamon.

Oller, J. 1979. Language Tests at School. London: Longman.

Rea, P. 1985. Language testing and the communicative language teaching curriculum. In Lee, Fok, Lord and Low (eds.)

Richards, J. 1985. Planning for proficiency. Prospect, 1, 2.

Ryle, G. 1949. The Concept of Mind. London; Hutchinson.

Savignon, S. and Berns, M. (eds.). 1984. Initiatives in Communicative Language Teaching. Reading Mass.: Addison-Wesley.